DS 613 Graduate Project Guidelines

Find a data table that is original and different from the data tables used in class, and in need of tidying and general transformation, and has quantitative as well as categorical variables:

* Generally, you are to apply procedures and techniques covered in class to analyze and process the data that we covered in class. R coding is preferred, but **you can also feature a different coding tool for Data Analysis, such as Python. You can also make use of a different IDE, such as Google Colab, Visual Studio, or Tableau.**
* Fully explain the data, providing insight regarding all variables of the data table
* Use and show R coding that presents your table as a data frame and a tibble
* Use additional R code covered in class to identify and characterize your data table
* Use tidyverse coding to import the data into R
* Use tidy R and dplyr functions to modify the data so that it is structured properly for better analysis and processing
* Use ggplot coding to generate plots that convey interesting and impactful facts and trends of selected categorical and quantitative variables.
* Show evidence of new R coding to process and analyze data in the table that we did not cover in class

Additional notes:

Try to find a table that is reasonably large and robust with at least 10 variables and at least 100 observational rows.